

QA
76
M94
SI-Ref

M U D P I E no. 22

Museum and University Data, Program and Information Exchange

SURVEY OF COMPUTER PROGRAM CATALOGS-IV

"NAVAL ORDNANCE LABORATORY LIBRARY OF COMPUTER PROGRAMS"

The U. S. Naval Ordnance Laboratory (NOL) has developed a subroutine library "to gather and to distribute computer subroutines and to prevent wasteful duplication. This system provides a method for programmers to obtain write-ups and decks for already existing routines which perform functions they want." While developed primarily for the use of their own staff, NOL will provide information to other workers upon request. As of May, 1971, they were grouping their programs under the following index headings (the number in parentheses indicates the total number of programs available in that class):

1. Utility and Systems (31)
2. Programmed Arithmetic (0)
3. Elementary Functions (3)
4. Polynomials, root finders, special functions, non-linear equations (25)
5. Differentiation, Integration, Differential Equations (11)
6. Interpolation, Approximation, Curve Fitting (14)
7. Matrices, Vectors, Simultaneous Linear Equations (10)
8. Statistical Analysis (20)
9. Operations Research and Optimization (1)
10. Information Processing (4)
11. Engineering Applications (0)
12. Miscellaneous (0)

Further information and/or a list of the available programs in each of the index categories can be obtained from: Leah E. Brown, Computer Applications Division, U.S. Naval Ordnance Laboratory, White Oak, Silver Spring MD 20910.--James A. Peters.

IDENTIFICATION OF BIOLOGICAL SPECIMENS BY COMPUTER

A three-year project on identification of biological specimens by computer began at the Botany School, Cambridge, on 1 June 1971. The project is funded by the Science Research Council, with Dr. S. M. Walters as principal investigator, assisted by Mr. R. J. Pankhurst and Ms. R. R. Cox. The project covers three main areas:

1. Improvement of existing computer programs for the generation of keys. A system for on-line keys will also be prepared.



2. Development of methods to recognize specimens by matching with descriptions in a data bank. This is intended to avoid some of the difficulties of using keys. One example which is being investigated is the microspecies of Rubus fruticosus in Great Britain.
3. Development of methods of data capture. Whatever method of identification is used, large volumes of accurate data must be obtained. Computer methods are being used to reduce the labor involved.

Anyone working on, or interested in, the above topics, who is not acquainted with our project is invited to make contact with us at the Botany School, Downing Street, Cambridge CB2 3EA England.
--S. M. Walters & R. J. Pankhurst.

[Editors note: Two reports on progress in this research program have been issued in mimeo form, as follows:

- Pankhurst, R. J. A method for data capture. Identification Project Report no. 1, 1971: 1-8.
Pankhurst, R. J. Identification by matching-Stage One (Rubus). Identification Project Report no. 2, 1971: 1-7.

Copies are available at the address above.--jap]

MEETING ON AUTOMATIC IDENTIFICATION

The Systematics Association will sponsor a meeting on automatic identification in September, 1973, at King's College. Only preliminary announcements have been issued to date, but further information will appear in MUDPIE as available. Interested people can contact R. J. Pankhurst at the address given above.

M U S T A R D

[Museum and University Storage and Retrieval of Data]

Beginning with this issue of MUDPIE all items dealing with data storage and retrieval will be collected together under this heading, and will, of course, be known as MUSTARD papers. In the last year a rather considerable body of interest has emerged in the storage and retrieval of data about specimens, chiefly centered on specimens in museum collections. With this issue many individuals with this interest have been added to the list of recipients of MUDPIE. We hope to expand the list even further. Perhaps this section of MUDPIE will eventually come to serve the community of natural scientists as does the excellent "Newsletter of Computer Archaeology" for the archaeological community.

Information on projects underway or contemplated is sought

for inclusion in this section. It is hoped that coverage at least on an occasional basis of all such projects will become possible, and that this current awareness type of information will be of value to MUDPIE readers. Comments on this new departure for MUDPIE and contributions to the section are welcome.

THE NATIONAL MUSEUM OF NATURAL HISTORY EDP PROGRAM

The National Museum of Natural History received modest funding from Congress in Fiscal Year 1971 to begin the entry of data about specimens in collections under its care. Data were entered on all new incoming specimens in the Department of Paleobiology plus selected collections in the Departments of Vertebrate Zoology, Invertebrate Zoology and Botany. The key premise underlying the program is that museums of natural history have a heavy responsibility to care for the data related to natural history specimens just as they do for the specimens themselves. Prior to the advent of the computer as a data manipulator and ultimate data receptacle, and prior to the development of sophisticated methods of data input, it was simply impossible to efficiently manipulate large quantities of data about individual specimens. It was recognized in the NMNH EDP program that in order better to fulfill this responsibility for care of information about collections it would be necessary to demonstrate that advantages could be obtained not only in the distant future but in the immediate environment of day-to-day museum operations. Consequently considerable effort was expended in eliminating or greatly reducing the amount of repetitive copying of data in the production of the documents needed in curating. Intensive effort is also being devoted to the production of intermediate level aids to better use of the collections such as computer-created and cross-indexed catalogs. The first of these catalogs, a catalog of Conodont type-specimens, was published in September 1971. In the future, catalogs on NMNH and other museum holdings of botanical type specimens will be forthcoming as will further catalogs or special collections in the other departments mentioned above.

Two more departments of the NMNH will be brought into the EDP program in Fiscal Year 1972. The Department of Entomology is preparing data for a revision of the authoritative Catalog of Hymenoptera last revised in 1967. The current revision, under the direction of Karl Krombein, will be typed on paper tape typewriters and read into the Smithsonian's Honeywell 1250 computer. The SELGEM data storage and retrieval system will be used to accept and process these data, and, after editing, correction and update runs, will produce a magnetic tape ready for the computer-driven Linotron at the Government Printing Office. A good deal of planning still remains to be done before the project is fully implemented, but there seem to be no insurmountable obstacles to carrying this project to completion. The immediate result of this type of data processing will be the production of a catalog



published in fully diversified type styles. The long-range product will be the creation of a data base which can be added to in the course of intervening years so that subsequent revisions of the catalog will not be as tedious to produce. Another obvious benefit will be the availability of the file for querying. All categories of information considered by the hymenopterists to be of significantly great interest to warrant separate identification will be so identified in the computer file and terms in these categories will be searchable.

Two files will be created in the Department of Mineral Sciences. One of these will be a computerized index to the very important petrologic collections in the care of the Department. Another file, containing a great variety of information on the volcanoes of the world, will also be built. Work on the Volcanic Activity File has already begun and work on the petrologic collection will soon begin.

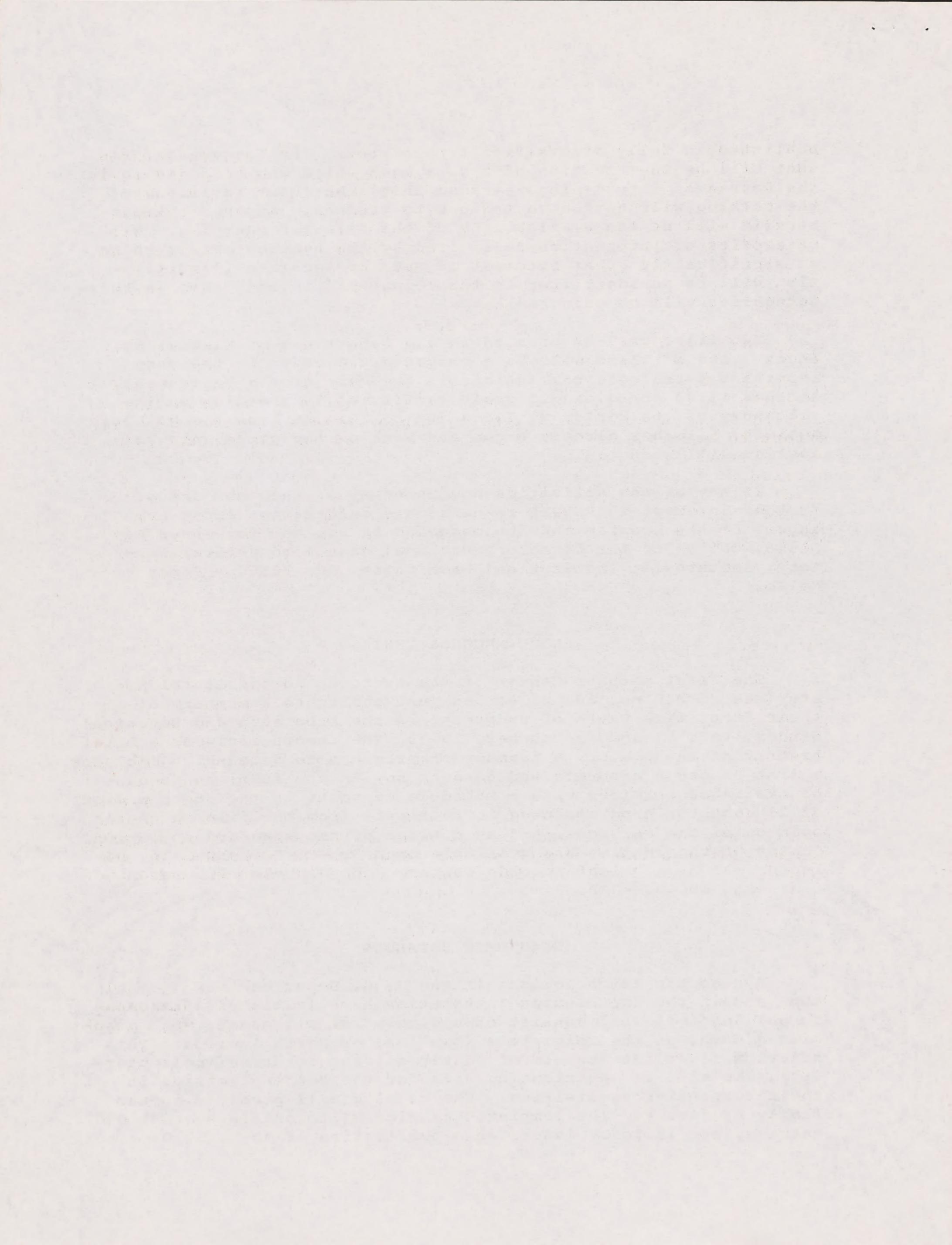
If any of the activities now underway at the NMNH are of further interest to MUDPIE readers more information about any aspect of the program can be obtained by writing to: James F. Mello, Office of the Director, National Museum of Natural History, Smithsonian Institution, Washington DC 20560.--James F. Mello.

PLANT RECORDS CENTER

The Plant Records Center of the American Horticultural Society (see MUDPIE no. 18, p. 4) has just published a summary of their first five years of progress, in the Arboretum and Botanical Garden, vol. 5, no. 4, October, 1971. The report includes a brief history of the P.R.C.; a summary of progress to date which includes a list of those arboreta and gardens now in the file; an account of available services with a schedule of costs to the user; a list of 21 computer programs used in retrieval from the file; a user's opinion of the services; and an outline of the standard operation. Copies of the leaflet are available from Dr. R. D. MacDonald, Director, A. H. S. Plant Records Center, John J. Tyler Arboretum, Lima PA 19060.--JAP.

TAXONOMIC CATALOGS

Systematic entomologists in the U. S. Department of Agriculture and in the Smithsonian Institution have initiated a taxonomic catalog that will consist of a listing of all names, with associated data, of the Coleoptera (beetles) of North America. This effort will enlist the aid of nearly all the concerned coleopterists, who will be contributing data for the beetle families in their respective specialties. The catalog will probably appear family by family. The completed catalog will contain over 40,000 entries, and is to be issued as a publication of the U.S. Dept. of



Agriculture. In evaluating this catalog for publication, the Publication Committee of the Agricultural Research Service, USDA, estimated that the total cost would be about \$100,000 if the catalog were to be published in a formal way. The Committee approved a pilot project in which all of the data would be committed initially to computer memory, with later programs to be prepared so that camera-ready copy could be computer-generated. Published versions of the catalog could then be printed by offset, and at the same time the data could be kept permanently in the computer file for additions and other updating procedures; for producing subsequently updated versions of the printed catalog; and for querying by remote terminals, batch processing, etc. The pilot project will produce the immediate objective of a printed catalog of one or two families to test cost-effectiveness, and a long range objective of testing maintenance and file query capabilities. The entire effort to automate this data has been placed in the hands of the Data Systems Application Division (DSA), Agricultural Research Service. Thus far, two beetle families have been completely cataloged--the Anobiidae by Richard White and the Dermestidae by John Kingsolver. The input-correction program has been completed and is ready for testing. The maintenance and query programs are yet to be written. Input will be by MT/ST tape; and at the present writing the first such tape, including possibly half of one of the families, is being produced.

A similar project has been planned by specialists in the Hymenoptera (bees, wasps, ants), consisting of a catalog of the Hymenoptera of North America, a new version of an older catalog and its supplements. This catalog will be published as a single volume rather than family by family, and currently plans are being made to produce it by computer-generated cold copy as in the case of the Coleoptera catalog. The printed versions of these two catalogs will follow different formats, dictated in part by the circumstances of the scientists who will generate the data. Even more recently, specialists in the Lepidoptera (butterflies and moths) have expressed their interest in these efforts. Although they are not free to mount a major cataloging effort due to other commitments, their data can be accommodated with only minor adjustments. Because the proposed system lends itself admirably to the storage of small bits and pieces of catalog information, the current research produced by the lepidopterists can be used as source material and adapted to the catalog mode.

Because the four major orders of insects (Diptera, Coleoptera, Lepidoptera, and Hymenoptera) are involved in this project, a means to initiate the stockpiling of current data for at least three-fourths of all insect species is within reach. Because insects make up four-fifths of all living animals, the way will be opened toward the establishment of an automated world inventory of living things.--(Abstracted from a statement by Richard H. Foote).

SEAMAK ZOOGAD SYSTEM

This system for the storage, retrieval and analysis of data concerning animals in zoos around the world has been designed and implemented by U. S. Seal and D. G. Makey, in collaboration with the American Association of Zoo Veterinarians. The system has been designed to store and retrieve records consisting of hematology, blood chemistry and physical condition. Since the system intends to deal with all of the mammalian species, a code for the taxonomic names of all such species was developed, consisting of a twelve digit alphanumeric code for each taxon, including subspecies. A zoogeographic code of four numeric digits uniquely codes all existing zoos in the world and also identifies geographic regions.

Two magnetic tapes are used, the first containing the system programs, mammalian taxonomic codes and zoogeographic codes, the second containing the physiological data records. The system includes two programs, one to record new records on tape, and one to recall records on the basis of six parameters, which are:

- A. Taxonomic code.
- B. Animal identification number.
- C. Animal location.
- D. Data source.
- E. Animal origin.
- F. A disease classification.

A preliminary summary of the system to date has been issued, which gives instructions on the format for data input, shows what output format is, lists the entire zoogeographic code (71 pages), which includes all or most of the world's zoos, also lists the entire mammalian species code (206 pages), and gives a preliminary numerical code to the orders of the remaining vertebrate classes. Anyone interested in further information on the system should write to Dr. U. S. Seal, V. A. Hospital, 54th St. and 48th Avenue South, Minneapolis MN 55417.--JAP.

STORAGE AND RETRIEVAL OF BIOLOGICAL INFORMATION

The texts of a series of lectures arranged by the Linnean Society of London and the Committee on Biological Information and presented in London from 5 October to 14 December 1970 on the subject above have been published in the Biological Journal of the Linnean Society, vol. 3, no. 3, Sept. 1971, taking up the entire issue. The following papers are included:

- Edwards, P.I. The general pattern of biological information.
Edwards, P.I. List of libraries in the field of pure and applied biology.
Onions, A.H.S. Culture collections of microfungi.
Lapage, S.P. Culture collections of bacteria.
Whitehead, P.J.P. Storage and retrieval of information in system-

- atic biology.
Edwards, P.I. Zoos and aquaria.
Edwards, P.I. Special zoological indexes.
Stearn, W.T. Sources of information about botanic gardens and herbaria.
Edwards, P.I. Special botanical indexes.
Perring, F. The Biological Records Centre--a data centre.
Eggins, H.O.W. The Biodeterioration Information Centre: specialized information centre.
Edwards, P.I. Information and data centres in the biological sciences.
Cutbill, J.L. New methods for handling biological information.
Abbot, M. Patents.
Edwards, P.I. Report and trade literature.
Edwards, P.I. Some botanical reference works.
Edwards, P.I. List of abstracting and indexing services in pure and applied biology.
Scrivenor, T. V. The Commonwealth Agricultural Bureaux abstracting and indexing services.
Dadd, M.N. The Zoological Record--current developments.
Meikle, R.D. The history of the Index Kewensis.

RECENT LITERATURE

- Bonham, C. D. A computer program for mapping ecological data. Science Series no. 9, Range Science Dept., Colorado State University, 1971: 1-30 plus 15 pp. of program.
Estabrook, G. F. Some information theoretic optimality criteria for general classification. Journal of the International Association for Mathematical Geology, vol. 3, 1971: 203-208.
Howarth, R. J. FORTRAN IV program for grey-level mapping of spatial data. Journal of the International Association for Mathematical Geology, vol. 3, 1971: 95-122.
Mosmann, C. J. Computing in higher education: successes and prospects--a report on the fall council meeting. EDUCOM Bulletin, vol. 6, 1971: 2-9.
Vickers, F. D. Painless teaching via computer: an information system for a large course. EDUCOM-Bulletin, vol. 6, 1971: 10-12.

February, 1972
Division of Reptiles and Amphibians
National Museum of Natural History
Washington, DC 20560

